CONFERENCIA INTERNACIONAL BIREDIAL-ISTEC '26
20-23 de octubre de 2026 Lima • PERÚ

UNR

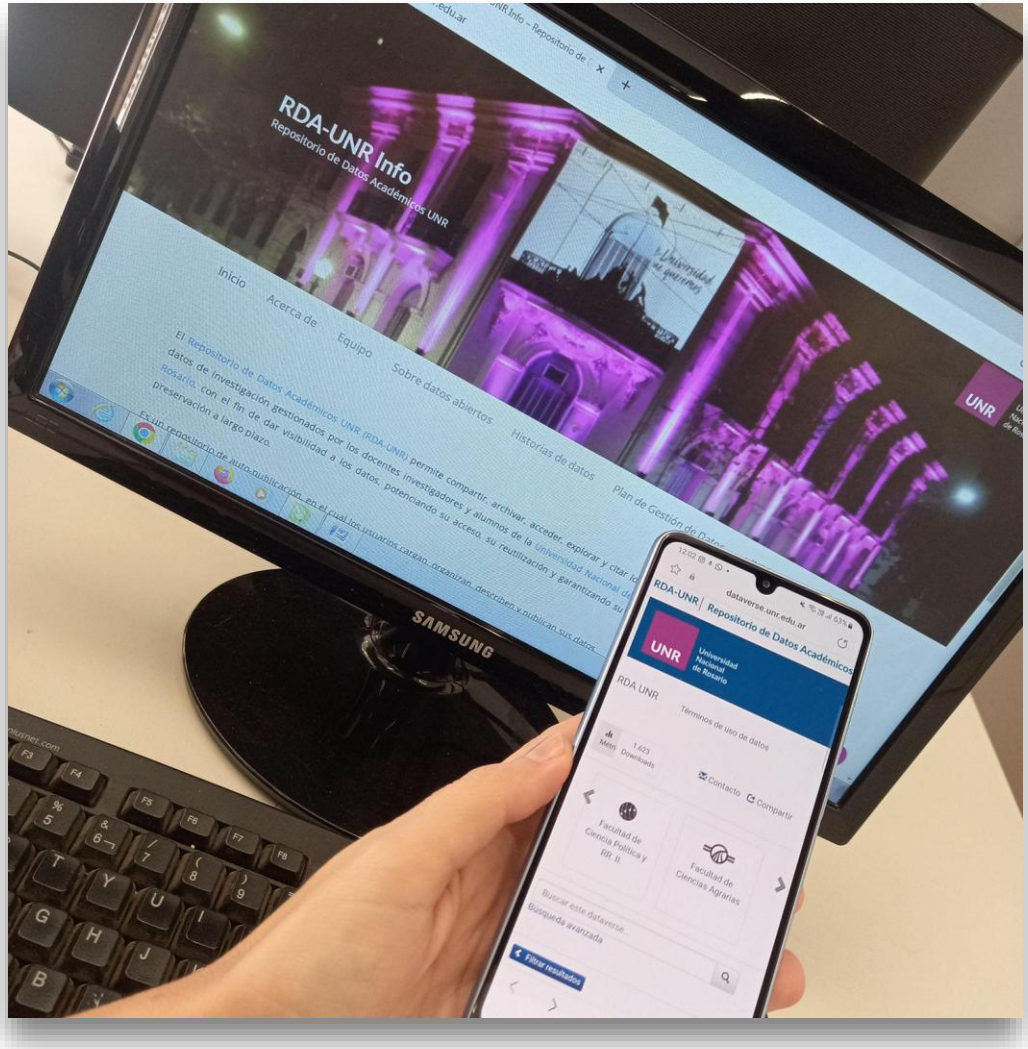
Publicar datos en abierto con calidad FAIR

Curación, documentación y visibilidad en el RDA-UNR

Dra. Paola Bongiovani
Universidad Nacional de Rosario, Argentina

Esta presentación se distribuye bajo una licencia Creative Commons 4.0 de Atribución

Acceso Abierto



El RDA-UNR en contexto

Desde dónde hablamos

El RDA-UNR en la Universidad Nacional de Rosario

Primer repositorio de datos en universidades argentinas · desde agosto de 2022

La UNR

- 87.000 estudiantes
- 8.100 profesores
- 3.500 investigadores
- 13 unidades académicas y 15 bibliotecas

Acceso abierto

- 2008: RepHipUNR
- 2016: Comité de Acceso Abierto y UGAA
- 2013-2016: Ley 26.899 y su reglamentación
- 2017-2018: políticas de acceso abierto
- 2020: Portal Revistas UNR
- 2021-2022: Repositorio de Datos Académicos

RDA-UNR

- Software Dataverse y DOIs vía DataCite.
- Auto Depósito con acompañamiento: consulta de curación y revisión antes de publicar.
- Equipo interdisciplinario: información, sistemas e investigación
- dataverse.unr.edu.ar

Profesionales de la información

Competencias que ya tenemos

- Organización de la información
- Aplicación de estándares y metadatos
- Consciencia sobre la importancia de la preservación
- Contacto con docentes e investigadores
- Formación de usuarios

Lo que se suma en el camino

- Ciclo de vida y curación de datos
- Identificadores persistentes y cita de datos
- Licencias y datos sensibles
- Conocer las prácticas de cada disciplina
- Trabajo interdisciplinario

Se aprende con la práctica y en equipo: no hace falta saber todo para empezar

LOS DATOS NO “HABLAN POR SÍ SOLOS”

Su interpretación depende del contexto en el que se produjeron y cómo se usan (Leonelli, 2016).

La creación de datos no es neutra: está atravesada por contextos, decisiones y negociaciones (Knorr Cetina, 2005).

- **Contextualizar los datos: metadatos y documentación**
- **Reconocer los límites y alcances de la reutilización**
- **Cuidar el significado de los datos cuando circulan**

Por eso un dataset se construye entre quien investiga y curadores, es un proceso interdisciplinario.



Acceso Abierto

Los datos y FAIR

Qué son, cómo se describen y por qué es importante acompañar

¿Qué son los datos de investigación?

Información recolectada, observada o generada durante el proceso de investigación que fundamenta las afirmaciones y sirve para validar los resultados de una investigación (RDA-UNR)

Observacionales

Audios de movimientos mandibulares en ganado lechero (Vanrell et al., 2022)

Teledetección

Imágenes Radarsat-2 en horticultura intensiva (Cotlier et al., 2022)

Encuesta

Factores que influyen en compartir datos en la UNR (Bongiovani, 2024)

Clínicos

515 episodios de neutropenia febril en Argentina (Parodi et al., 2022)

Computacionales

Modelación matemática hidrológica de la cuenca del A° Pavón (Stenta et al., 2022)

Documentales

Cuaderno de Demandas del Juzgado de Paz de San Telmo, 1836 (Gastellu, 2025)

Ejemplos publicados en el RDA-UNR

Principios FAIR, con ejemplos concretos



Findable (Encontrables)

Identificadores persistentes y metadatos ricos, indexados en buscadores. Título y descripción centrados en los datos.



Accessible (Accesibles)

Protocolos estándar. Los metadatos están visibles aunque los archivos estén restringidos.



Interoperable (Interoperables)

Dublin Core, DDI, DataCite. Vocabularios controlados como Tesauro UNESCO, DeCS/MeSH o AGROVOC.



Reusable (Reutilizables)

Licencia clara, procedencia (cómo y quién) y estándares de la disciplina. Documentación de los datos.

Herramientas de evaluación automática, como F-UJI, ayudan a revisar, pero no reemplazan la mirada de la persona curadora.

Abierto no es igual a FAIR

❌ Abierto pero NO FAIR

- Archivo Excel compartido en Google Drive sin descripción.
- Sin DOI, sin metadatos, sin licencia explícita.
- Nadie puede encontrarlo ni sabe qué contiene.
- Imposible reproducir sin contactar al autor.
- El propio equipo de investigadores puede necesitar contexto.

✅ Abierto y FAIR

- Dataset depositado en RDA-UNR con DOI y otros identificadores.
- Metadatos descriptivos completos (título, autores, metodología, etc.).
- Licencia explícita.
- Documentación.
- README con descripción de variables y unidades.

FAIR y CARE: los datos y las personas

FAIR mira los datos; CARE mira a las comunidades de las que provienen. Son complementarios.

C — Beneficio colectivo

Los datos deben beneficiar a los pueblos de los que provienen.

A — Autoridad para controlar

Los pueblos tienen derecho a decidir sobre sus datos.

R — Responsabilidad

Quien usa los datos rinde cuentas sobre cómo los usa.

E — Ética

Los derechos y el bienestar de los pueblos, en el centro de todo el ciclo de los datos.

La persona curadora pregunta:

- ¿Los datos provienen de pueblos o comunidades indígenas?
- ¿Participaron en la decisión de publicarlos?
- ¿Qué condiciones de uso pidieron?

Carroll, S. R. et al. (2020). The CARE Principles for Indigenous Data Governance. Data Science Journal, 19, 43. https://doi.org/10.5334/dsj-2020-043
· Global Indigenous Data Alliance (GIDA)

¿Por qué importa acompañar? Evidencia de la UNR

Encuesta a 412 investigadores/as de UNR y CONICET (57,5% UNR · 31,8% UNR-CONICET · 10,7% CIUNR)
Bongiovani, P. C. (2024). Tesis Doctoral, Universidad Carlos III de Madrid. https://hdl.handle.net/10016/44153

H2

✓ **Respaldada**

Altruismo Académico

Refleja las creencias de los investigadores sobre los beneficios de compartir datos para la ciencia, la sociedad y otros investigadores.

H5

✓ **Respaldada**

Conocimientos y Habilidades

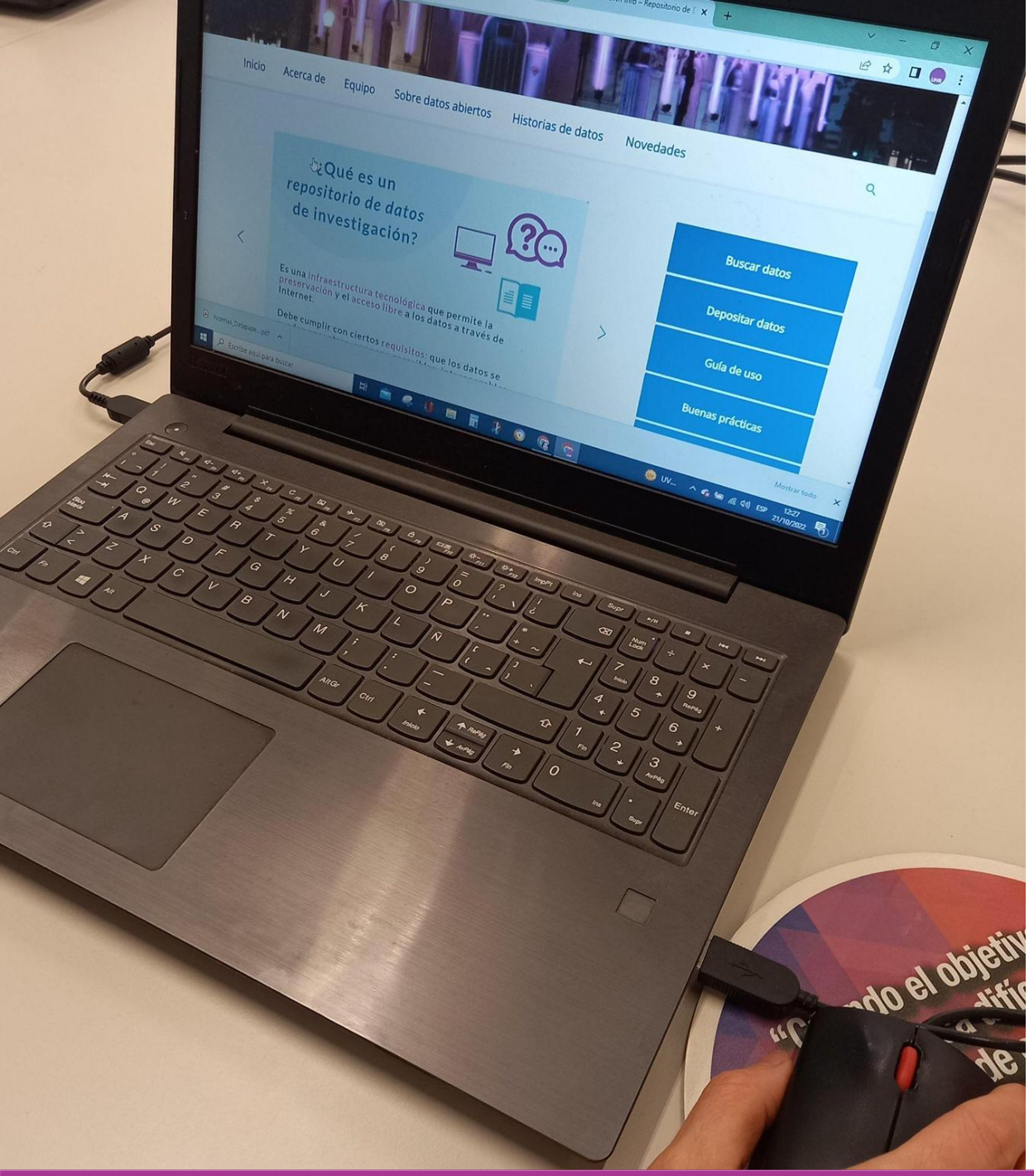
Los investigadores que manifestaron mayor confianza en sus competencias de gestión de datos expresaron una mayor intención de compartir datos.

H9

✓ **Respaldada**

Disponibilidad de Infraestructura y Repositorios

Los investigadores que percibían que existían repositorios accesibles estaban más dispuestos a compartir sus datos.



Acceso Abierto

La curación de datos

Rol y herramientas

¿QUÉ ES LA CURACIÓN DE DATOS?

Gestión activa y valoración de los datos a lo largo de su ciclo de vida, para hacerlos útiles y accesibles para futuros usos.



En el RDA-UNR

- **Profesionales de la información trabajan con los investigadores para publicar sus datos**
- **Curadores especializados por área o generalistas**
- **Consulta de curación y revisión antes de publicar**

Los investigadores conocen sus datos: deciden qué datos funcionan como evidencia (Leonelli, 2015).

Hace falta saber preguntar

Investigador/a

- Conoce sus datos mejor que nadie
- Sabe cómo y por qué se generaron
- Decide qué publicar
- Lo obvio para el equipo puede no serlo para otros

Curador/a

- Lee el dataset como lo leería alguien externo, para entenderlo
- Traduce FAIR, metadatos y licencias a un lenguaje claro
- Pregunta lo que falta y propone mejoras
- Ayuda a que el dataset cuente una historia

No saber es parte de la herramienta: es la mirada de quien llegará a los datos sin conocer el proyecto.

En la consulta de curación preguntamos...

CONTENIDO

- ¿Cuál es el objetivo y qué rol cumplen los datos?
- ¿Cómo se recolectaron? ¿Con qué instrumentos?
- Unidad de análisis, período y cobertura

ESTRUCTURA

- ¿Qué contiene cada archivo?
- ¿Cómo se relacionan los archivos?
- ¿Hay diccionario de datos o libro de códigos?

CONTEXTO

- ¿Qué procesamiento tuvieron los datos?
- ¿Hay una publicación asociada?
- ¿Quién responde consultas a futuro?

ASPECTOS LEGALES O ÉTICOS

- ¿Hay datos de personas o sensibles?
- ¿Qué dice el consentimiento informado?
- ¿Hay datos de terceros o restricciones?

Guía RDA-UNR para describir los datos: preguntas

¿Sobre qué es?

Contenido

Tema, tipo de datos, cobertura temporal y geográfica, volumen y formato.

¿Cómo se obtuvo?

Procedencia

Contexto, metodología e instrumentos: el recorrido de los datos.

¿Cuál es su función?

Propósito

Qué evidencia aportan y para qué pregunta se generaron.

¿Por qué es único?

Valor

Qué lo distingue: lugar, período, población o método.

¿A quién le sirve?

Audiencia

Disciplinas, docentes, organismos u otros públicos.

¿Cómo reutilizarlo?

Reutilización

Usos posibles más allá del proyecto que lo produjo.

Los obstáculos más frecuentes y cómo los abordamos

Obstáculos

- El resumen copia el del artículo
- El título es el del artículo o la tesis
- Poca familiaridad con los metadatos
- Falta README o diccionario de datos

Soluciones en el RDA-UNR

- Guía para redactar la descripción
- Ejemplos de buenas descripciones
- Validación de archivos y formatos
- Documentación a dos niveles: proyecto y datos
- Lista de revisión para la curaduría

La descripción se centra en los datos, no en el artículo: se escribe para quien llega desde un buscador.

¿QUÉ DATOS PUBLICAR?

Los que sostienen los resultados y los que tienen valor para ser reutilizados.

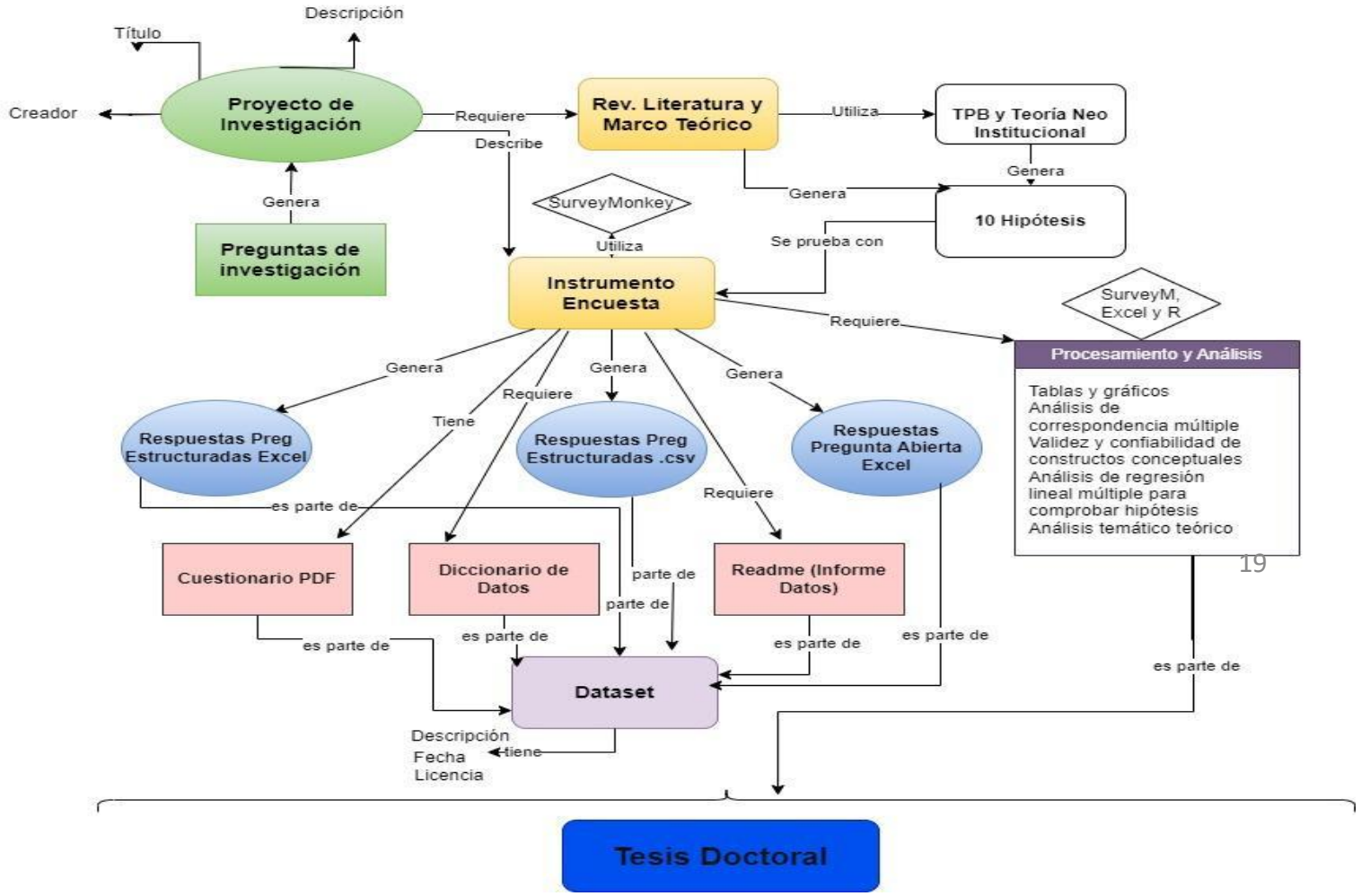
Modelar el flujo de la investigación ayuda a decidir qué datos son la evidencia.

- **¿Se pueden leer y usar con lo que se publica?**
- **¿Alcanzan las descripciones para que otra persona los entienda?**
- **¿Hay aspectos legales o éticos que impidan publicarlos?**

"Si hacés una tomografía tenés 500 mil cortes... ¿qué imagen se sube?"

Investigadora en ciencias médicas, capacitación en la UNR

Modelado del flujo de la investigación



Bongiovani, Paola Carolina, 2024, "Datos de encuesta sobre factores individuales e institucionales que influyen en la intención de compartir datos de investigación en la Universidad Nacional de Rosario", https://doi.org/10.57715/UNR/HLH331 , RDA UNR, V1,

Acceso Abierto

Publicar, cuidar, difundir

Ética, licencias y visibilidad



Ética y licencias: resolver, preguntar o derivar

"Tan abierto como sea posible, tan cerrado como sea necesario" UNESCO (2021)

Resolver

- Asignar una licencia clara (CC BY, CC0)
- Revisar que no haya datos personales directos
- Documentar las condiciones de acceso

Preguntar al equipo

- ¿El consentimiento contempla la publicación?
- ¿Hay datos de terceros?
- ¿Hay variables que, combinadas, identifican a alguien?

Derivar

- Datos sensibles según la ley de cada país (salud, origen étnico…)
- Comité de ética o área legal
- Acceso restringido o embargo, si corresponde

Visibilidad, citación e impacto

DOI y cita de datos

Cada dataset recibe un DOI y una cita estándar. Se encuentra en DataCite Commons y Google Dataset Search.

ORCID y filiación normalizada

Vinculan datos, publicaciones y personas, y atribuyen la producción a la institución.

≈ 25 % más citas

Vincular artículos con datos depositados en repositorios se asocia con más citas (Colavizza et al., 2020).

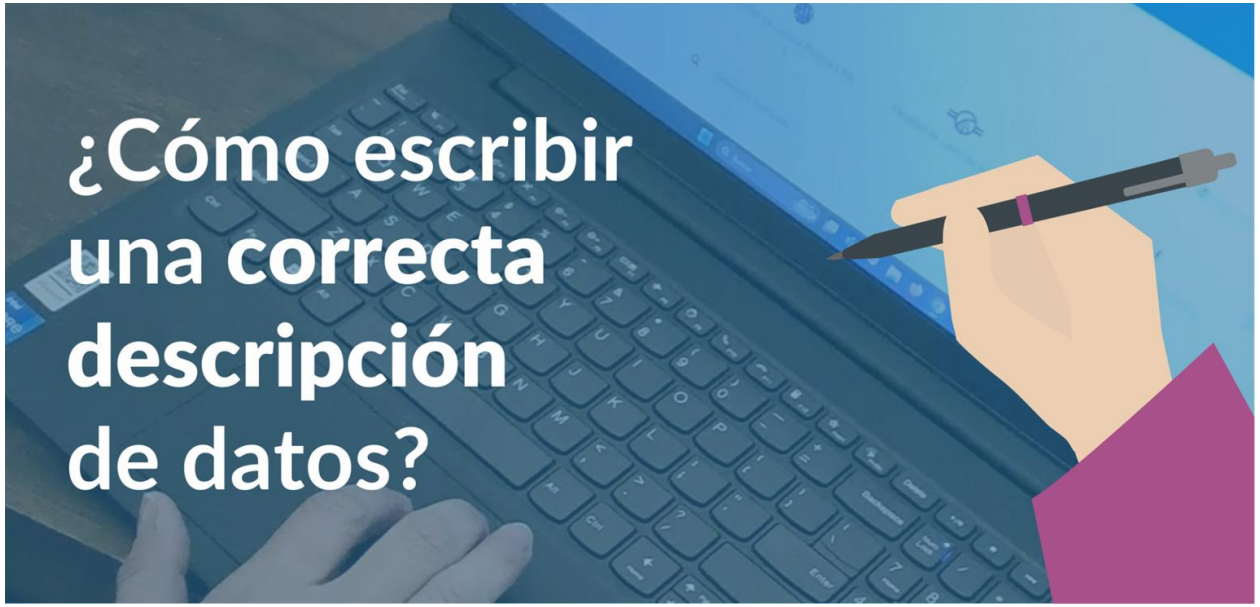
Historias de Datos

Videos y relatos en los que quienes investigan cuentan sus datos: acercan los datasets a otros públicos.

Acceso Abierto

Ahora, los casos

Seis preguntas para describir · lista de revisión · plantilla de README



CURATION: una guía para no olvidarnos de nada

Comprender

¿Qué datos hay? ¿Se entienden?

- **C** Comprobar
- **U** Entender

Hoy, en los casos

Mejorar con el equipo

Mejoro lo que puedo, pregunto lo que no sé

- **R** Recomendar mejoras
- **A** Ampliar o enriquecer

Hoy, en los casos

Preparar para publicar

Archivos listos, con DOI y licencia

- **T** Transformar formatos
- **I** Incluir identificadores y licencia

Cerrar

Evalúo y dejo constancia

- **O** Optimizar para FAIR
- **N** Anotar lo realizado

No hay un modelo único de curación: cada institución define el nivel que puede ofrecer.

CURATION (Portage Network; Cooper et al., 2021), adaptado de CURATE(D) de la Data Curation Network (Johnston et al., 2018). Traducción al español: curso de LA Referencia.

Revisamos entre todos con la Lista de revisión del RDA-UNR

Lista elaborada a partir de la práctica de curación, para asegurar la calidad del dataset y su reutilización.

- Ubicación y colección
- Título
- Autores
- Descripción de los datos
- Palabras clave
- Publicación relacionada
- Tipos de datos
- Software
- Formato de datos
- Revisión del contenido
- README y diccionario
- Datos sensibles

La descripción se organiza en seis apartados: introducción, tabla de contenidos, metodología, organización, calidad y valor.

Lista de revisión
RDA-UNR

Ficha de primera revisión
para seguir los casos

Recursos: estamos acompañados en el camino

Curso LA Referencia

lareferencia.info

Introducción a la gestión y curación de datos en Dataverse: cuatro módulos con casos.

Buenas prácticas

dataverse-info.unr.edu.ar

Guía para describir datos, formatos recomendados y lista de revisión.

Data Curation Network

datacurationnetwork.org

Guías (primers) de curación por tipo de archivo.

Guía Portage

doi.org/10.7191/jeslib.2021.1201

Base del marco CURATION para Dataverse (Cooper et al., 2021).

re3data.org

re3data.org

Directorio mundial de repositorios de datos.

Pop Ups

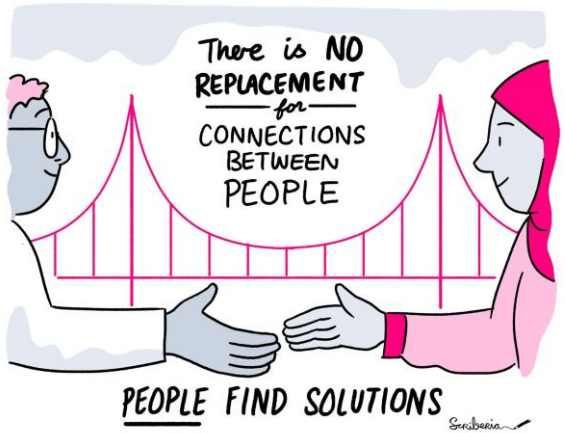
repositoriodatos@unr.edu.ar

Intercambio de experiencias y trabajo en red.



No se puede reemplazar la conexión humana

Encontramos soluciones juntos



Equipo del RDA-UNR en Biredial-ISTEC 2024

BIBLIOGRAFÍA

Bongiovani, P. C. (2024). Factores individuales e institucionales que influyen en la intención de compartir datos de investigación en repositorios: el caso de la Universidad Nacional de Rosario [Tesis doctoral, Universidad Carlos III de Madrid]. https://hdl.handle.net/10016/44153

Bongiovani, P. C., Díaz Pacífico, F. y Freán, P. (2025). Datos FAIR en Argentina: desarrollo y desafíos del Repositorio de Datos Académicos RDA-UNR. Información, Cultura y Sociedad, (52), 175-202. https://doi.org/10.34096/ics.i52.16501

Colavizza, G., Hrynaszkiewicz, I., Staden, I., Whitaker, K. y McGillivray, B. (2020). The citation advantage of linking publications to research data. PLOS ONE, 15(4), e0230416. https://doi.org/10.1371/journal.pone.0230416

Cooper, A. et al. (2021). Creating guidance for Canadian Dataverse curators: Portage Network's Dataverse Curation Guide. Journal of eScience Librarianship, 10(3), e1201. https://doi.org/10.7191/jeslib.2021.1201

Knorr Cetina, K. (2005). La fabricación del conocimiento. Universidad Nacional de Quilmes.

Lee, D. J. y Stvilia, B. (2017). Practices of research data curation in institutional repositories. PLOS ONE, 12(3), e0173987. https://doi.org/10.1371/journal.pone.0173987

Leonelli, S. (2015). What counts as scientific data? A relational framework. Philosophy of Science, 82(5), 810-821. https://doi.org/10.1086/684083

Leonelli, S. (2016). Data-centric biology: A philosophical study. University of Chicago Press.

Schmidt, B. y Shearer, K. (2016). Librarians' competencies profile for research data management. COAR.

Wilkinson, M. D. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018. https://doi.org/10.1038/sdata.2016.18

¡Muchas gracias!

dataverse-info.unr.edu.ar

repositoriodatos@unr.edu.ar
pbongio@unr.edu.ar

CC BY

UNR